WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 01:08 PM GMT

PDB ID	:	3SZE
Title	:	Crystal structure of the passenger domain of the E. coli autotransporter EspP
Authors	:	Khan, S.; Mian, H.S.; Sandercock, L.E.; Battaile, K.P.; Lam, R.; Chirgadze, N.Y.; Pai, E.F.
Deposited on	:	2011-07-18
Resolution	:	2.50 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
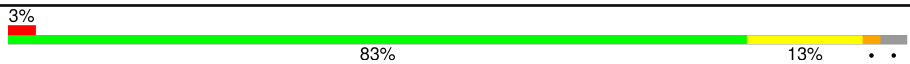*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
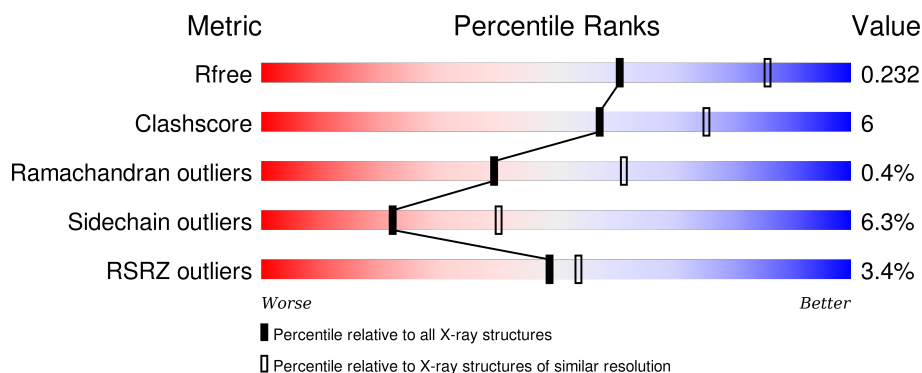
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	968	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7576 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Serine protease espP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	943	7252	4512	1251	1477	12	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	263	ALA	SER	ENGINEERED MUTATION	UNP Q7BSW5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	324	324	324	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Serine protease espP
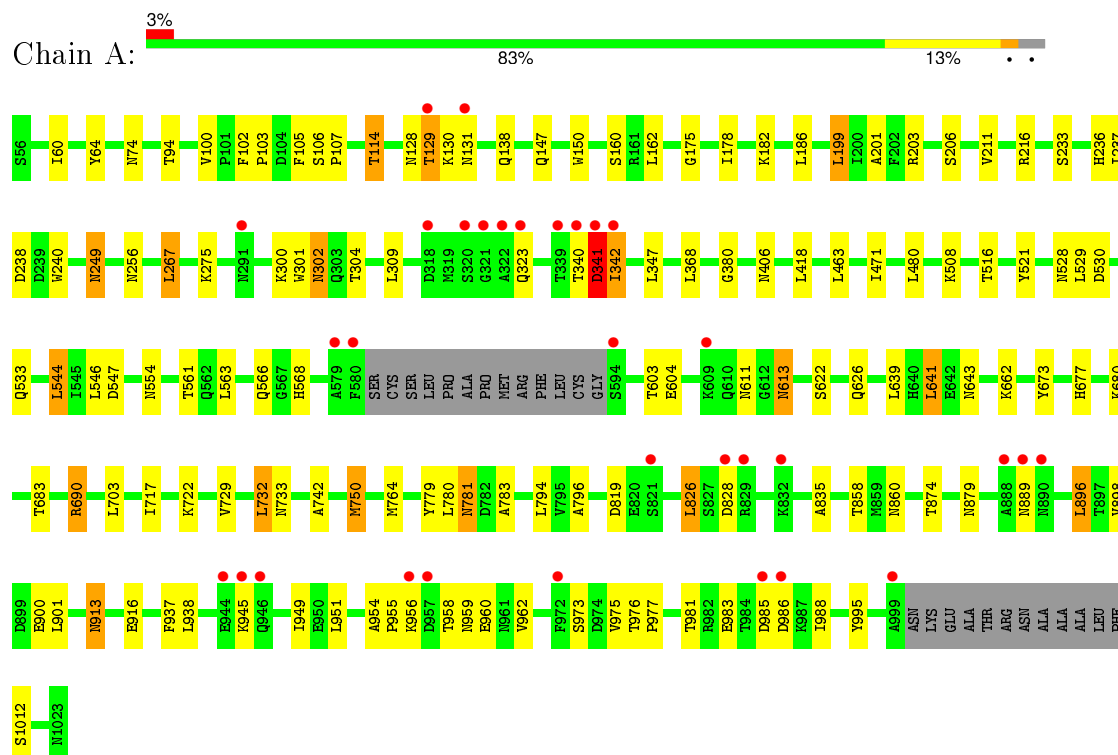
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	88.34Å 88.34Å 311.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 29.55 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.50) 99.8 (29.55-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.47 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.188 , 0.231 0.192 , 0.232	Depositor DCC
R_{free} test set	2193 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	30.2	Xtriage
Anisotropy	0.182	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 33.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Outliers	0 of 43649 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7576	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z| > 5$	RMSZ	#$|Z| > 5$
1	A	0.41	0/7383	0.57	0/9998

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7252	0	6999	85	0
2	A	324	0	0	3	0
All	All	7576	0	6999	85	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (85) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:342:ILE:HD13	1:A:347:LEU:HD21	1.37	1.06
1:A:341:ASP:HA	1:A:342:ILE:CG2	1.95	0.97
1:A:114:THR:HG21	1:A:138:GLN:NE2	1.80	0.96
1:A:74:ASN:HD21	1:A:105:PHE:H	1.25	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:342:ILE:CD1	1:A:347:LEU:HD21	2.08	0.83
1:A:341:ASP:HA	1:A:342:ILE:HG22	1.61	0.79
1:A:114:THR:HG21	1:A:138:GLN:HE21	1.47	0.78
1:A:983:GLU:HG2	1:A:988:ILE:HG22	1.66	0.77
1:A:613:ASN:HD22	1:A:613:ASN:N	1.83	0.77
1:A:613:ASN:HD22	1:A:613:ASN:H	1.35	0.75
1:A:341:ASP:CA	1:A:342:ILE:HG22	2.17	0.74
1:A:341:ASP:HA	1:A:342:ILE:HG23	1.68	0.72
1:A:114:THR:CG2	1:A:138:GLN:NE2	2.54	0.71
1:A:201:ALA:HB1	1:A:267:LEU:HD22	1.72	0.69
1:A:732:LEU:HD22	1:A:733:ASN:OD1	1.92	0.69
1:A:528:ASN:HD22	1:A:554:ASN:HD21	1.43	0.65
1:A:566:GLN:HE22	1:A:673:TYR:H	1.44	0.63
1:A:147:GLN:NE2	1:A:160:SER:OG	2.31	0.63
1:A:60:ILE:HD11	1:A:603:THR:HG22	1.81	0.62
1:A:342:ILE:HD13	1:A:347:LEU:CD2	2.23	0.61
1:A:954:ALA:HB1	1:A:955:PRO:HD2	1.86	0.57
1:A:703:LEU:HD23	1:A:722:LYS:HB2	1.87	0.56
1:A:896:LEU:HD22	1:A:898:VAL:HG23	1.87	0.56
1:A:249:ASN:N	1:A:249:ASN:HD22	2.04	0.55
1:A:199:LEU:HD13	1:A:237:ILE:HB	1.89	0.55
1:A:860:ASN:HD21	1:A:879:ASN:HD22	1.55	0.54
1:A:60:ILE:HD11	1:A:603:THR:CG2	2.38	0.54
1:A:340:THR:O	1:A:341:ASP:HB3	2.06	0.54
1:A:508:LYS:HD3	1:A:530:ASP:OD2	2.08	0.54
1:A:913:ASN:HD21	1:A:916:GLU:H	1.56	0.53
1:A:956:LYS:HA	1:A:988:ILE:HG23	1.90	0.53
1:A:463:LEU:HD13	1:A:471:ILE:HD11	1.91	0.52
1:A:211:VAL:HG21	1:A:603:THR:HG21	1.91	0.52
1:A:147:GLN:HE21	1:A:160:SER:CB	2.22	0.52
1:A:983:GLU:CG	1:A:988:ILE:HG22	2.38	0.51
1:A:566:GLN:NE2	1:A:673:TYR:H	2.08	0.50
1:A:463:LEU:CD1	1:A:471:ILE:HD11	2.41	0.50
1:A:781:ASN:ND2	1:A:783:ALA:H	2.09	0.50
1:A:233:SER:HB3	2:A:1414:HOH:O	2.11	0.49
1:A:201:ALA:HB1	1:A:267:LEU:CD2	2.42	0.49
1:A:717:ILE:HG22	1:A:742:ALA:HB1	1.94	0.49
1:A:74:ASN:ND2	1:A:105:PHE:H	2.04	0.48
1:A:566:GLN:HE22	1:A:673:TYR:N	2.10	0.47
1:A:547:ASP:OD1	1:A:568:HIS:HD2	1.97	0.47
1:A:764:MET:CE	1:A:794:LEU:HD11	2.44	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:SER:HB2	1:A:256:ASN:HB3	1.97	0.47
1:A:544:LEU:HD13	1:A:563:LEU:CD2	2.44	0.47
1:A:236:HIS:HD2	1:A:238:ASP:OD2	1.98	0.46
1:A:182:LYS:HZ2	1:A:240:TRP:H	1.63	0.46
1:A:976:THR:HG23	1:A:977:PRO:HD2	1.97	0.45
1:A:129:THR:HG22	1:A:131:ASN:H	1.81	0.45
1:A:764:MET:O	1:A:796:ALA:HA	2.16	0.45
1:A:302:ASN:C	1:A:302:ASN:HD22	2.20	0.44
1:A:106:SER:N	1:A:107:PRO:CD	2.81	0.44
1:A:613:ASN:N	1:A:613:ASN:ND2	2.55	0.44
1:A:639:LEU:HG	1:A:641:LEU:HD13	2.00	0.43
1:A:826:LEU:HD21	1:A:835:ALA:HB1	2.00	0.43
1:A:102:PHE:CG	1:A:103:PRO:HD2	2.53	0.43
1:A:975:VAL:HG13	1:A:995:TYR:HB2	2.00	0.43
1:A:643:ASN:HA	1:A:662:LYS:O	2.18	0.43
1:A:683:THR:CG2	1:A:690:ARG:HD2	2.49	0.43
1:A:781:ASN:HD22	1:A:783:ALA:H	1.66	0.43
1:A:521:TYR:HA	2:A:1105:HOH:O	2.19	0.43
1:A:677:HIS:CD2	1:A:690:ARG:HG3	2.53	0.42
1:A:342:ILE:C	1:A:342:ILE:HD12	2.39	0.42
1:A:937:PHE:CE1	1:A:975:VAL:HG12	2.53	0.42
1:A:611:ASN:HB2	1:A:613:ASN:ND2	2.35	0.42
1:A:729:VAL:HB	1:A:750:MET:HG2	2.02	0.42
1:A:128:ASN:O	1:A:129:THR:CB	2.68	0.41
1:A:626:GLN:NE2	2:A:1230:HOH:O	2.52	0.41
1:A:533:GLN:NE2	1:A:561:THR:OG1	2.53	0.41
1:A:958:THR:HB	1:A:988:ILE:HD11	2.03	0.41
1:A:341:ASP:N	1:A:342:ILE:HG22	2.35	0.41
1:A:949:ILE:O	1:A:951:LEU:HD13	2.21	0.41
1:A:874:THR:HG22	1:A:900:GLU:HB2	2.01	0.41
1:A:129:THR:HG22	1:A:130:LYS:N	2.36	0.41
1:A:732:LEU:HD22	1:A:733:ASN:CG	2.41	0.41
1:A:622:SER:HA	1:A:626:GLN:HE22	1.86	0.41
1:A:64:TYR:HA	1:A:604:GLU:OE2	2.21	0.41
1:A:175:GLY:HA3	1:A:301:TRP:CE2	2.56	0.41
1:A:178:ILE:HG21	1:A:300:LYS:HB3	2.03	0.41
1:A:913:ASN:C	1:A:913:ASN:HD22	2.25	0.40
1:A:683:THR:HG22	1:A:690:ARG:HD2	2.03	0.40
1:A:959:ASN:O	1:A:962:VAL:HG23	2.21	0.40
1:A:750:MET:HB2	1:A:779:TYR:HD2	1.87	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	938/968 (97%)	887 (95%)	47 (5%)	4 (0%)	39	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	341	ASP
1	A	129	THR
1	A	323	GLN
1	A	380	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	791/810 (98%)	741 (94%)	50 (6%)	22	40

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	94	THR
1	A	100	VAL
1	A	114	THR
1	A	150	TRP
1	A	162	LEU
1	A	186	LEU
1	A	199	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	203	ARG
1	A	206	SER
1	A	216	ARG
1	A	249	ASN
1	A	267	LEU
1	A	275	LYS
1	A	302	ASN
1	A	304	THR
1	A	309	LEU
1	A	341	ASP
1	A	342	ILE
1	A	368	LEU
1	A	406	ASN
1	A	418	LEU
1	A	480	LEU
1	A	516	THR
1	A	529	LEU
1	A	544	LEU
1	A	546	LEU
1	A	613	ASN
1	A	641	LEU
1	A	680	LYS
1	A	690	ARG
1	A	732	LEU
1	A	750	MET
1	A	780	LEU
1	A	781	ASN
1	A	819	ASP
1	A	826	LEU
1	A	828	ASP
1	A	858	THR
1	A	889	ASN
1	A	896	LEU
1	A	901	LEU
1	A	913	ASN
1	A	938	LEU
1	A	945	LYS
1	A	960	GLU
1	A	973	SER
1	A	981	THR
1	A	985	ASP
1	A	986	ASP

Continued on next page...

Mol	Chain	Res	Type
1	A	1012	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (32) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	ASN
1	A	138	GLN
1	A	142	ASN
1	A	147	GLN
1	A	151	ASN
1	A	236	HIS
1	A	249	ASN
1	A	250	GLN
1	A	288	ASN
1	A	294	ASN
1	A	302	ASN
1	A	401	ASN
1	A	406	ASN
1	A	515	ASN
1	A	526	ASN
1	A	528	ASN
1	A	533	GLN
1	A	566	GLN
1	A	568	HIS
1	A	613	ASN
1	A	626	GLN
1	A	643	ASN
1	A	653	ASN
1	A	691	GLN
1	A	699	GLN
1	A	740	ASN
1	A	781	ASN
1	A	798	ASN
1	A	818	HIS
1	A	860	ASN
1	A	913	ASN
1	A	931	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	943/968 (97%)	-0.15	32 (3%)	49	54	17, 29, 55, 77	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	985	ASP	5.0
1	A	986	ASP	5.0
1	A	888	ALA	4.7
1	A	972	PHE	4.6
1	A	889	ASN	4.2
1	A	944	GLU	4.2
1	A	821	SER	4.1
1	A	342	ILE	4.1
1	A	828	ASP	3.7
1	A	131	ASN	3.5
1	A	341	ASP	3.3
1	A	291	ASN	3.1
1	A	321	GLY	3.1
1	A	339	THR	2.9
1	A	129	THR	2.9
1	A	340	THR	2.8
1	A	579	ALA	2.8
1	A	832	LYS	2.8
1	A	956	LYS	2.7
1	A	945	LYS	2.7
1	A	594	SER	2.6
1	A	323	GLN	2.6
1	A	890	ASN	2.6
1	A	322	ALA	2.4
1	A	320	SER	2.3
1	A	946	GLN	2.2
1	A	609	LYS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	957	ASP	2.1
1	A	999	ALA	2.1
1	A	829	ARG	2.0
1	A	318	ASP	2.0
1	A	580	PHE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.